

SEQUENCE LISTING

<110> THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
GILL, Gordon N.
YEO, Michele
LIN, Patrick S.
DAHMUS, Michael E.

<120> PHOSPHATASE REGULATION OF NUCLEIC ACID TRANSCRIPTION

<130> 00015-041US1

<140> US/10/552,298
<141> 2005-09-30

<150> US 60/459,786
<151> 2003-04-01

<160> 69

<170> PatentIn version 3.5

<210> 1
<211> 783
<212> DNA
<213> Homo sapiens

<400> 1
atggacagct cggccgtcat tactcagatc agcaaggagg aggctcgaaa cccgctgcgg 60
ggcaaagggtg accagaagtc agcagttcc cagaagcccc gaagccgggg catcctccac 120
tcactcttct gctgtgtctg ccggatgtat ggggaggccc tgcctgctca cagcggggcg 180
ccctgcttg tggaggagaa tggccatc cctaagaccc cagtccaata cctgctccct 240
gaggccaagg cccaggactc agacaagatc tgcgtggtca tcgacctgga cgagaccctg 300
gtgcacagct cttcaagcc agtgaacaac gcggacttca tcattccctgt ggagattgtat 360
ggggtgttcc accaggtcta cgtgttgaag cgtcctcatg tggatgagtt cctgcagcga 420
atgggcgagc tcttgaatg tgtgtgttc actgcttagcc tcgccaagta cgcagaccca 480
gtagctgacc tgctggacaa atggggggcc ttccggggcc ggctgtttcg agagtccctgc 540
gtcttccacc gggggaaacta cgtgaaggac ctgagccgggt tgggtcgaga cctgcggcgg 600
gtgctcatcc tggacaattc acctgcctcc tatgtcttcc atccagacaa tgctgtaccg 660
gtggcctcgt gtgttgcacaa catgagtgac acagagctcc acgacctcct ccccttcttc 720
gagcaactca gccgtgtgga cgacgtgtac tcagtgctca ggcagccacg gccaggagc 780
tag 783

<210> 2
<211> 260
<212> PRT
<213> Homo sapiens

<400> 2

Met Asp Ser Ser Ala Val Ile Thr Gln Ile Ser Lys Glu Glu Ala Arg
1 5 10 15

Gly Pro Leu Arg Gly Lys Gly Asp Gln Lys Ser Ala Ala Ser Gln Lys
20 25 30

Pro Arg Ser Arg Gly Ile Leu His Ser Leu Phe Cys Cys Val Cys Arg
35 40 45

Asp Asp Gly Glu Ala Leu Pro Ala His Ser Gly Ala Pro Leu Leu Val
50 55 60

Glu Glu Asn Gly Ala Ile Pro Lys Thr Pro Val Gln Tyr Leu Leu Pro
65 70 75 80

Glu Ala Lys Ala Gln Asp Ser Asp Lys Ile Cys Val Val Ile Asp Leu
85 90 95

Asp Glu Thr Leu Val His Ser Ser Phe Lys Pro Val Asn Asn Ala Asp
100 105 110

Phe Ile Ile Pro Val Glu Ile Asp Gly Val Val His Gln Val Tyr Val
115 120 125

Leu Lys Arg Pro His Val Asp Glu Phe Leu Gln Arg Met Gly Glu Leu
130 135 140

Phe Glu Cys Val Leu Phe Thr Ala Ser Leu Ala Lys Tyr Ala Asp Pro
145 150 155 160

Val Ala Asp Leu Leu Asp Lys Trp Gly Ala Phe Arg Ala Arg Leu Phe
165 170 175

Arg Glu Ser Cys Val Phe His Arg Gly Asn Tyr Val Lys Asp Leu Ser
180 185 190

Arg Leu Gly Arg Asp Leu Arg Arg Val Leu Ile Leu Asp Asn Ser Pro
195 200 205

Ala Ser Tyr Val Phe His Pro Asp Asn Ala Val Pro Val Ala Ser Trp
210 215 220

Phe Asp Asn Met Ser Asp Thr Glu Leu His Asp Leu Leu Pro Phe Phe
225 230 235 240

Glu Gln Leu Ser Arg Val Asp Asp Val Tyr Ser Val Leu Arg Gln Pro
245 250 255

Arg Pro Gly Ser
260

<210> 3
<211> 852
<212> DNA
<213> Homo sapiens

<400> 3
atggaacacg gctccatcat cacccaggcg cgagggaaag acgccttgtt gctcaccaag 60
caaggcctgg tctccaagtc ctctcctaag aagcctcggt gacgtaacat cttcaaggcc 120
cttttctgct gtttcgcgc ccagcatgtt ggccagtcaa gttcctccac tgagctcgct 180
gcgtataagg aggaagcaaa caccattgct aagtccggatc tgctccagtg tctccagttac 240
cagttctacc agatcccagg gacctgcctg ctcccagagg tgacagagga agatcaagga 300
aggatctgtg tggtcattga cctcgatgaa acccttgtgc atagctcctt taagccaatc 360
aacaatgctg acttcatagt gcctatagag attgagggga ccactcacca ggtgtatgtg 420
ctcaagaggc cttatgtgga ttagttcctg agacgcatgg gggactctt tgaatgtgtt 480
ctcttcactg ccagcctggc caagtatgcc gaccctgtga cagacctgct ggaccgggt 540
ggggtgttcc gggcccgccct attccgtgag tcttgcgtgt tccaccaggg ctgctacgtc 600
aaggacctca gccgcctggg gagggacctg agaaagaccc tcatcctgga caactcgcc 660
gcttcttaca tattccaccc cgagaatgca gtgcctgtgc agtcctggtt ttagtgcacatg 720
gcagacactg agttgctgaa cctgatccca atcttgagg agctgagcgg agcagaggac 780
gtctacacca gccttggggc agctgcgggc cccttagcct gcccgttcc caagcgacgg 840
ccatcccagt ag 852

<210> 4
<211> 283
<212> PRT
<213> Homo sapiens

<400> 4

Met Glu His Gly Ser Ile Ile Thr Gln Ala Arg Arg Glu Asp Ala Leu
1 5 10 15

Val Leu Thr Lys Gln Gly Leu Val Ser Lys Ser Ser Pro Lys Lys Pro
20 25 30

Arg Gly Arg Asn Ile Phe Lys Ala Leu Phe Cys Cys Phe Arg Ala Gln
35 40 45

His Val Gly Gln Ser Ser Ser Thr Glu Leu Ala Ala Tyr Lys Glu
50 55 60

Glu Ala Asn Thr Ile Ala Lys Ser Asp Leu Leu Gln Cys Leu Gln Tyr
65 70 75 80

Gln Phe Tyr Gln Ile Pro Gly Thr Cys Leu Leu Pro Glu Val Thr Glu
85 90 95

Glu Asp Gln Gly Arg Ile Cys Val Val Ile Asp Leu Asp Glu Thr Leu
100 105 110

Val His Ser Ser Phe Lys Pro Ile Asn Asn Ala Asp Phe Ile Val Pro
115 120 125

Ile Glu Ile Glu Gly Thr Thr His Gln Val Tyr Val Leu Lys Arg Pro
130 135 140

Tyr Val Asp Glu Phe Leu Arg Arg Met Gly Glu Leu Phe Glu Cys Val
145 150 155 160

Leu Phe Thr Ala Ser Leu Ala Lys Tyr Ala Asp Pro Val Thr Asp Leu
165 170 175

Leu Asp Arg Cys Gly Val Phe Arg Ala Arg Leu Phe Arg Glu Ser Cys
180 185 190

Val Phe His Gln Gly Cys Tyr Val Lys Asp Leu Ser Arg Leu Gly Arg
195 200 205

Asp Leu Arg Lys Thr Leu Ile Leu Asp Asn Ser Pro Ala Ser Tyr Ile
210 215 220

Phe His Pro Glu Asn Ala Val Pro Val Gln Ser Trp Phe Asp Asp Met
225 230 235 240

Ala Asp Thr Glu Leu Leu Asn Leu Ile Pro Ile Phe Glu Glu Leu Ser
245 250 255

Gly Ala Glu Asp Val Tyr Thr Ser Leu Gly Ala Ala Ala Gly Pro Leu
260 265 270

Ala Cys Pro Ala Ser Lys Arg Arg Pro Ser Gln
275 280

<210> 5
<211> 798
<212> DNA
<213> Homo sapiens

<400> 5
atggacggcc cgccccatcat cacccagggtg accaacccta aggaggacga gggccggttg 60
ccgggcgcgg gcgagaaagc ctcccagtgc aacgtcagct taaagaagca gaggagccgc 120
agcatcctta gtccttctt ctgctgcttc cgtgattaca atgtggaggc ccctccaccc 180
agcagccca gtgtgcttcc gccactggtg gaggagaatg gtgggcttca gaagccacca 240
gctaagtacc ttcttccaga ggtgacggtg cttgactatg gaaagaaatg tgtggtcatt 300
gatttagatg aaacatttgt gcacagttcg tttaagccta ttagtaatgc tgattttatt 360
gttccggttg aatcgatgg aactatacat caggtgtatg tgctgaagcg gccacatgtg 420
gacgagttcc tccagaggat gggcagctt tttgaatgtg tgctcttac tgccagctt 480
gccaagtatg cagaccctgt ggctgacctc ctagaccgct ggggtgtgtt ccggggccgg 540
ctcttcagag aatcatgtgt tttcatcgt gggactacg tgaaggacct gagtcgcctt 600
gggcgggagc tgagcaaagt gatcattgtt gacaattccc ctgcctcata catcttccat 660
cctgagaatg cagtcctgt gcagtcctgg ttcgatgaca tgacggacac ggagctgctg 720
gacctcatcc cttctttga gggcctgagc cgggaggacg acgtgtacag catgctgcac 780
agactctgca ataggtag 798

<210> 6
<211> 265
<212> PRT
<213> Homo sapiens

<400> 6

Met Asp Gly Pro Ala Ile Ile Thr Gln Val Thr Asn Pro Lys Glu Asp
1 5 10 15

Glu Gly Arg Leu Pro Gly Ala Gly Glu Lys Ala Ser Gln Cys Asn Val
20 25 30

Ser Leu Lys Lys Gln Arg Ser Arg Ser Ile Leu Ser Ser Phe Phe Cys
35 40 45

Cys Phe Arg Asp Tyr Asn Val Glu Ala Pro Pro Pro Ser Ser Pro Ser
50 55 60

| | | | | | | | | | | | | | | | |
|------------|--------------|------------|------------|------------|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Leu | Pro | Pro | Leu | Val | Glu | Glu | Asn | Gly | Gly | Leu | Gln | Lys | Pro | Pro |
| 65 | | | | | 70 | | | | | | 75 | | | | 80 |
| | | | | | | | | | | | | | | | |
| Ala | Lys | Tyr | Leu | Leu | Pro | Glu | Val | Thr | Val | Leu | Asp | Tyr | Gly | Lys | Lys |
| | | | | | 85 | | | | 90 | | | | 95 | | |
| | | | | | | | | | | | | | | | |
| Cys | Val | Val | Ile | Asp | Leu | Asp | Glu | Thr | Leu | Val | His | Ser | Ser | Phe | Lys |
| | | | | | 100 | | | 105 | | | | | | 110 | |
| | | | | | | | | | | | | | | | |
| Pro | Ile | Ser | Asn | Ala | Asp | Phe | Ile | Val | Pro | Val | Glu | Ile | Asp | Gly | Thr |
| | | | | | 115 | | | 120 | | | | | 125 | | |
| | | | | | | | | | | | | | | | |
| Ile | His | Gln | Val | Tyr | Val | Leu | Lys | Arg | Pro | His | Val | Asp | Glu | Phe | Leu |
| | | | | | 130 | | | 135 | | | | 140 | | | |
| | | | | | | | | | | | | | | | |
| Gln | Arg | Met | Gly | Gln | Leu | Phe | Glu | Cys | Val | Leu | Phe | Thr | Ala | Ser | Leu |
| | | | | | 145 | | | 150 | | | 155 | | | 160 | |
| | | | | | | | | | | | | | | | |
| Ala | Lys | Tyr | Ala | Asp | Pro | Val | Ala | Asp | Leu | Leu | Asp | Arg | Trp | Gly | Val |
| | | | | | 165 | | | 170 | | | | | 175 | | |
| | | | | | | | | | | | | | | | |
| Phe | Arg | Ala | Arg | Leu | Phe | Arg | Glu | Ser | Cys | Val | Phe | His | Arg | Gly | Asn |
| | | | | | 180 | | | 185 | | | | | 190 | | |
| | | | | | | | | | | | | | | | |
| Tyr | Val | Lys | Asp | Leu | Ser | Arg | Leu | Gly | Arg | Glu | Leu | Ser | Lys | Val | Ile |
| | | | | | 195 | | | 200 | | | | 205 | | | |
| | | | | | | | | | | | | | | | |
| Ile | Val | Asp | Asn | Ser | Pro | Ala | Ser | Tyr | Ile | Phe | His | Pro | Glu | Asn | Ala |
| | | | | | 210 | | | 215 | | | 220 | | | | |
| | | | | | | | | | | | | | | | |
| Val | Pro | Val | Gln | Ser | Trp | Phe | Asp | Asp | Met | Thr | Asp | Thr | Glu | Leu | Leu |
| | | | | | 225 | | | 230 | | | 235 | | | 240 | |
| | | | | | | | | | | | | | | | |
| Asp | Leu | Ile | Pro | Phe | Phe | Glu | Gly | Leu | Ser | Arg | Glu | Asp | Asp | Val | Tyr |
| | | | | | 245 | | | 250 | | | | | 255 | | |
| | | | | | | | | | | | | | | | |
| Ser | Met | Leu | His | Arg | Leu | Cys | Asn | Arg | | | | | | | |
| | | | | | 260 | | | 265 | | | | | | | |
| | | | | | | | | | | | | | | | |
| <210> | 7 | | | | | | | | | | | | | | |
| <211> | 642 | | | | | | | | | | | | | | |
| <212> | DNA | | | | | | | | | | | | | | |
| <213> | Homo sapiens | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
| <400> | 7 | | | | | | | | | | | | | | |
| atgatgggaa | gcccctgcct | gctcacagcg | gggcgcccct | gcttgtggag | gagaatggcg | | | | | | | | | | 60 |

| | | | | | | |
|------------|-------------|------------|------------|------------|------------|-----|
| ccatccctaa | ggcagacccc | agtccaatac | ctgctccctg | aggccaaggc | ccaggactca | 120 |
| gacaagatct | gcgtggtcat | cgacctggac | gagaccctgg | tgcacagctc | cttcaagcca | 180 |
| gtgaacaacg | cggacttcat | catccctgtg | gagattgatg | gggtggtcca | ccaggtctac | 240 |
| gtgttgaagc | gtcctcacgt | ggatgagttc | ctgcagcgaa | tggcgagct | ctttgaatgt | 300 |
| tgctgttca | ctgcttagcct | cgccaagtac | gcagacccag | tagctgacct | gctggacaaa | 360 |
| tggggggcct | tccggggcccg | gctgtttcga | gagtcctgcg | tcttccaccg | ggggactac | 420 |
| gtgaaggacc | tgagccgggtt | gggtcgagac | ctgcggcggg | tgctcatcct | ggacaattca | 480 |
| cctgcctcct | atgtcttcca | tccagacaat | gctgtaccgg | tggcctcgtg | gtttgacaac | 540 |
| atgagtgaca | cagagctcca | cgacccctc | cccttcttcg | agcaactca | ccgtgtggac | 600 |
| gacgtgtact | cagtgctcag | gcagccacgg | ccagggagct | ag | | 642 |

<210> 8
 <211> 213
 <212> PRT
 <213> Homo sapiens

<400> 8

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Met | Gly | Arg | Pro | Cys | Leu | Leu | Thr | Ala | Gly | Arg | Pro | Cys | Leu | Trp |
| 1 | | | | | 5 | | | | 10 | | | | | 15 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Arg | Met | Ala | Pro | Ser | Leu | Arg | Gln | Thr | Pro | Val | Gln | Tyr | Leu | Leu |
| | | 20 | | | | | 25 | | | | | 30 | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pro | Glu | Ala | Lys | Ala | Gln | Asp | Ser | Asp | Lys | Ile | Cys | Val | Val | Ile | Asp |
| | | 35 | | | | 40 | | | | | 45 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Asp | Glu | Thr | Leu | Val | His | Ser | Ser | Phe | Lys | Pro | Val | Asn | Asn | Ala |
| | | 50 | | | | 55 | | | | 60 | | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Phe | Ile | Ile | Pro | Val | Glu | Ile | Asp | Gly | Val | Val | His | Gln | Val | Tyr |
| | | 65 | | | 70 | | | 75 | | | | 80 | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Leu | Lys | Arg | Pro | His | Val | Asp | Glu | Phe | Leu | Gln | Arg | Met | Gly | Glu |
| | | | | | 85 | | | 90 | | | 95 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Phe | Glu | Cys | Val | Leu | Phe | Thr | Ala | Ser | Leu | Ala | Lys | Tyr | Ala | Asp |
| | | | | | 100 | | | 105 | | | 110 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pro | Val | Ala | Asp | Leu | Leu | Asp | Lys | Trp | Gly | Ala | Phe | Arg | Ala | Arg | Leu |
| | | | | | | 115 | | 120 | | | 125 | | | | |

Phe Arg Glu Ser Cys Val Phe His Arg Gly Asn Tyr Val Lys Asp Leu

130

135

140

Ser Arg Leu Gly Arg Asp Leu Arg Arg Val Leu Ile Leu Asp Asn Ser
145 150 155 160

Pro Ala Ser Tyr Val Phe His Pro Asp Asn Ala Val Pro Val Ala Ser
165 170 175

Trp Phe Asp Asn Met Ser Asp Thr Glu Leu His Asp Leu Leu Pro Phe
180 185 190

Phe Glu Gln Leu Ser Arg Val Asp Asp Val Tyr Ser Val Leu Arg Gln
195 200 205

Pro Arg Pro Gly Ser
210

<210> 9
<211> 783
<212> DNA
<213> Drosophila

<400> 9
atggacagct cggccgtcat tactcagatc agcaaggagg aggctcgaaa cccgctgcgg 60
ggcaaaggta accagaagtc agcagttcc cagaagcccc gaagccgggg catcctccac 120
tcactcttct gctgtgtctg ccggatgtat ggggaggccc tgcctgctca cagcggggcg 180
ccctgcttg tggaggagaa tggcccatc cctaagaccc cagtccaata cctgctccct 240
gaggccaagg cccaggactc agacaagatc tgcgtgtca tcgarctgaa cgagaccctg 300
gtgcacagct cttcaagcc agtgaacaac gcggacttca tcattccctgt ggagattgtat 360
ggggtgttcc accaggtcta cgtgttgaag cgtcctcatg tggatgagtt cctgcagcga 420
atgggcgagc tcttgaatg tgtgctgttc actgcttagcc tcgccaagta cgcaaaaa 480
gtagctgacc tgctggacaa atggggggcc ttccggggcc ggctgtttcg agagtcctgc 540
gtcttccacc gggggacta cgtgaaggac ctgagccgggt tgggtcgaga cctgcggcgg 600
gtgctcatcc tggacaattc acctgcctcc tatgtcttcc atccagacaa tgctgtaccg 660
gtggcctcgt gtgttgcacaa catgagtgac acagagctcc acgacccctt ccccttcttc 720
gagcaactca gccgtgtgga cgacgtgtac tcagtgctca ggcagccacg gccaggagc 780
tag 783

<210> 10
<211> 260
<212> PRT

<213> Drosophila

<400> 10

Met Asp Ser Ser Ala Val Ile Thr Gln Ile Ser Lys Glu Glu Ala Arg
1 5 10 15

Gly Pro Leu Arg Gly Lys Gly Asp Gln Lys Ser Ala Ala Ser Gln Lys
20 25 30

Pro Arg Ser Arg Gly Ile Leu His Ser Leu Phe Cys Cys Val Cys Arg
35 40 45

Asp Asp Gly Glu Ala Leu Pro Ala His Ser Gly Ala Pro Leu Leu Val
50 55 60

Glu Glu Asn Gly Ala Ile Pro Lys Thr Pro Val Gln Tyr Leu Leu Pro
65 70 75 80

Glu Ala Lys Ala Gln Asp Ser Asp Lys Ile Cys Val Val Ile Glu Leu
85 90 95

Asn Glu Thr Leu Val His Ser Ser Phe Lys Pro Val Asn Asn Ala Asp
100 105 110

Phe Ile Ile Pro Val Glu Ile Asp Gly Val Val His Gln Val Tyr Val
115 120 125

Leu Lys Arg Pro His Val Asp Glu Phe Leu Gln Arg Met Gly Glu Leu
130 135 140

Phe Glu Cys Val Leu Phe Thr Ala Ser Leu Ala Lys Tyr Ala Asp Pro
145 150 155 160

Val Ala Asp Leu Leu Asp Lys Trp Gly Ala Phe Arg Ala Arg Leu Phe
165 170 175

Arg Glu Ser Cys Val Phe His Arg Gly Asn Tyr Val Lys Asp Leu Ser
180 185 190

Arg Leu Gly Arg Asp Leu Arg Arg Val Leu Ile Leu Asp Asn Ser Pro
195 200 205

Ala Ser Tyr Val Phe His Pro Asp Asn Ala Val Pro Val Ala Ser Trp
210 215 220

Phe Asp Asn Met Ser Asp Thr Glu Leu His Asp Leu Leu Pro Phe Phe

225 230 235 240

Glu Gln Leu Ser Arg Val Asp Asp Val Tyr Ser Val Leu Arg Gln Pro
245 250 255

Arg Pro Gly Ser
260

<210> 11
<211> 642
<212> DNA
<213> Drosophila

<400> 11
atgatgggga ggcctgcct gtcacagcg gggcgccccct gcttgtggag gagaatggcg 60
ccatccctaa ggcagacccc agtccaatac ctgctccctg aggccaaggc ccaggactca 120
gacaagatct gcgtggtcat cgarcgtgaac gagaccctgg tgcacagctc cttcaagcca 180
gtgaacaacg cggacttcat catccctgtg gagattgtat gggtggtcca ccaggtctac 240
gtgttgaagc gtcctcacgt ggatgagttc ctgcagcgaa tggcgagct ctttgaatgt 300
gtgctgttca ctgcttagcct cgccaagtac gcagacccag tagctgacct gctggacaaa 360
tggggggccct tccgggcccc gctgttcga gagtcctgcg tcttccaccg ggggaactac 420
gtgaaggacc tgagccgggtt gggtcgagac ctgcggcggg tgctcatcct ggacaattca 480
cctgcctcct atgtttcca tccagacaat gctgtaccgg tggcctcg 540
atgagtgaca cagagctcca cgaccctcctc cccttcttcg agcaactcag ccgtgtggac 600
gacgtgtact cagtgctca gcaagccacgg ccagggagct ag 642

<210> 12
<211> 213
<212> PRT
<213> Drosophila

<400> 12

Met Met Gly Arg Pro Cys Leu Leu Thr Ala Gly Arg Pro Cys Leu Trp
1 5 10 15

Arg Arg Met Ala Pro Ser Leu Arg Gln Thr Pro Val Gln Tyr Leu Leu
20 25 30

Pro Glu Ala Lys Ala Gln Asp Ser Asp Lys Ile Cys Val Val Ile Glu
35 40 45

Leu Asn Glu Thr Leu Val His Ser Ser Phe Lys Pro Val Asn Asn Ala
50 55 60

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Phe | Ile | Ile | Pro | Val | Glu | Ile | Asp | Gly | Val | Val | His | Gln | Val | Tyr |
| 65 | | | | 70 | | | | 75 | | | | 80 | | | |
| Val | Leu | Lys | Arg | Pro | His | Val | Asp | Glu | Phe | Leu | Gln | Arg | Met | Gly | Glu |
| | | | | 85 | | | | 90 | | | | 95 | | | |
| Leu | Phe | Glu | Cys | Val | Leu | Phe | Thr | Ala | Ser | Leu | Ala | Lys | Tyr | Ala | Asp |
| | | | | 100 | | | | 105 | | | | 110 | | | |
| Pro | Val | Ala | Asp | Leu | Leu | Asp | Lys | Trp | Gly | Ala | Phe | Arg | Ala | Arg | Leu |
| | | | | 115 | | | | 120 | | | | 125 | | | |
| Phe | Arg | Glu | Ser | Cys | Val | Phe | His | Arg | Gly | Asn | Tyr | Val | Lys | Asp | Leu |
| | | | | 130 | | | | 135 | | | | 140 | | | |
| Ser | Arg | Leu | Gly | Arg | Asp | Leu | Arg | Arg | Val | Leu | Ile | Leu | Asp | Asn | Ser |
| | | | | 145 | | | | 150 | | | | 155 | | | 160 |
| Pro | Ala | Ser | Tyr | Val | Phe | His | Pro | Asp | Asn | Ala | Val | Pro | Val | Ala | Ser |
| | | | | 165 | | | | 170 | | | | 175 | | | |
| Trp | Phe | Asp | Asn | Met | Ser | Asp | Thr | Glu | Leu | His | Asp | Leu | Leu | Pro | Phe |
| | | | | 180 | | | | 185 | | | | 190 | | | |
| Phe | Glu | Gln | Leu | Ser | Arg | Val | Asp | Asp | Val | Tyr | Ser | Val | Leu | Arg | Gln |
| | | | | 195 | | | | 200 | | | | 205 | | | |
| Pro | Arg | Pro | Gly | Ser | | | | | | | | | | | |
| | | | | 210 | | | | | | | | | | | |

| | | | | | | | | | | | | | | | |
|------------|------------|------------|------------|-------------|------------|--|--|--|--|--|--|--|--|--|-----|
| <210> | 13 | | | | | | | | | | | | | | |
| <211> | 7020 | | | | | | | | | | | | | | |
| <212> | DNA | | | | | | | | | | | | | | |
| <213> | Drosophila | | | | | | | | | | | | | | |
| <400> | 13 | | | | | | | | | | | | | | |
| ctggagcgcg | gcaggaaccc | ggccggcccc | gcctcccagt | ccgccttagcc | gcccgggtcc | | | | | | | | | | 60 |
| cagaagtggc | gaaagccgca | gccgagtcca | ggtcacgccc | aagccgttgc | ccttttaagg | | | | | | | | | | 120 |
| gggagccttg | aaacggcgcc | tgggttccat | gtttgcatcc | gcctcgccgg | aaggaaactc | | | | | | | | | | 180 |
| catgttgtaa | caaagtttcc | tccgcgcccc | ctccctcccc | ctccccctta | gaacctggct | | | | | | | | | | 240 |
| cccctccct | ccggagctcg | cggggatccc | tccctcccac | ccctccccctc | ccccccgcgc | | | | | | | | | | 300 |
| cccgattccg | gccccagccg | ggggggaggc | cgggcgcccc | ggccagagtc | cggccggagc | | | | | | | | | | 360 |
| ggagcgcgcc | cggccccatg | gacagctcg | ccgtcattac | tcaagatcagc | aaggaggagg | | | | | | | | | | 420 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ctcgggggccc | gctgcggggc | aaaggtaccg | gggctgcggg | gagggggccg | aagccggggc | 480 |
| gccgtggag | gagagaaggg | gccggatct | tccccagggg | agccgcccgc | gccgccccgg | 540 |
| gcggccgcct | tagctgtgcc | cgaagctccc | agccc gagag | ggagcaggga | gagagttga | 600 |
| actcagagga | ggctcagaga | cgcgggcccgg | ggcctggcgc | cttggggcg | ctcctgtccg | 660 |
| ctcgaggtga | gaaaaactgag | gcaggaatag | agagggaaact | ctttcggggg | tttcctggca | 720 |
| ggcattgcgt | ggtgcattggg | cgcggggccca | ccattggcgc | caatggggct | gtgagatggg | 780 |
| ggagctgagg | agggcgccta | tgggcacccc | gtgagactc | cgcggccaccc | cccaccccca | 840 |
| cccccccggg | ctgcggtccg | gtagggtctt | gggagggggc | gccgagggtga | cagcaggctg | 900 |
| gggaggcttg | gagggatctc | ccgccaacac | acagctacgt | tccccacaaa | cttcgcgtca | 960 |
| cgcgtggagg | cgcggacccc | ctcggaggca | cagagaggac | ggccggcact | tccaagagtc | 1020 |
| gcttggcgcc | cgcggggaga | gtcgtgcgcc | tagtgggcac | gcaccacccc | gcaaagcctc | 1080 |
| gcgcggccga | cgaggctgcg | tccccagcg | tggctgggcc | gggggtggggg | ggtctgtctt | 1140 |
| ctcctttcc | ccgtgtggac | ctcaggatct | ggacgctgcc | cccaggctcg | cccacccctcg | 1200 |
| cctgggtctg | gctgccccgg | aactgagggc | aaggtggaaa | ggctagttgc | agggggccgg | 1260 |
| aggggggtgg | gttgggaggg | gtatctgtca | atcaggctgc | tgggctccag | gtcggaggtc | 1320 |
| tgggcggggc | agggcaaaca | gatggccact | ggacactggc | cccaggccgc | gggactgcac | 1380 |
| ccctgcctct | gggcccagcc | gcagtgagga | cttcgtaccc | acgggggtgg | agaggatgg | 1440 |
| gggaggggcag | gggtggactg | ccctgggtcc | caggccctgg | ctgtcctgag | caggggtgct | 1500 |
| caggttaaggt | ggggtcagga | ggcaccccaa | tggggctgat | cagcagcagt | catggaggct | 1560 |
| gtgagaggca | gggagagagc | accccaggac | ctccttctcc | aggccacgca | ctccctatgt | 1620 |
| gggcgcctta | atacctgcta | gacctatttgc | tctggagct | gcaggagcct | tggagttgat | 1680 |
| tgtggagccc | tgacaggggc | gtttcagaga | aagtcaaggag | ctgccttcgt | gtgtctggat | 1740 |
| gaagggggcca | cgcaagatc | ctcctggccc | aggggttcac | acctgggcac | acatgcagga | 1800 |
| ttctgcaggc | cagtgtgcac | cgagcctcca | acttgtgcct | ccctacttca | ggtgaccaga | 1860 |
| agtcaagcgc | ttcccagaag | ccccgaagcc | ggggcatcct | ccactcactc | ttctgctgt | 1920 |
| tctgccggga | tgtatggggag | gccctgcctg | ctcacagcgg | ggcgccccctg | cttgtggagg | 1980 |
| agaatggcgc | catccctaag | gtgcgtgggg | gccaggtggg | gccacggggg | cacctggact | 2040 |
| cagtcttcag | ggcttttaggg | gaaggggctc | ctgactgagc | ttttcaggat | ggacttgcag | 2100 |
| acctgaaagt | gcagagtagg | agggtggcag | cctccctgc | caggccctgc | ccactgtggg | 2160 |
| gaaactgaat | tctccctcat | aagtggaaagc | tttttctac | cttggtttt | agagaggtct | 2220 |

| | | | | | | |
|-------------|------------|-------------|-------------|-------------|-------------|------|
| caaagagcca | agaggcctac | ccaagcccta | gagctggcag | gggcaaagct | gggaaggggg | 2280 |
| aagtatctgt | tcctggggcc | tggggttcct | ctggagacgg | ctagggggag | aagcctgcgt | 2340 |
| gggaggaagg | accaggcccc | gagagaggca | ccccagccag | ccccggccctc | cctacagcag | 2400 |
| accccagtcc | aatacctgct | ccctgaggcc | aaggcccagg | actcagacaa | gatctgcgtg | 2460 |
| gtcatcgacc | tggacgagac | cctggtgcac | agtccttca | aggtgggccc | tgctcaacag | 2520 |
| ccctcagccc | gggtctcggg | gggcattcccc | caccctggcc | tgggagggag | gtgtgtgctg | 2580 |
| gaccatgc | cctggggctc | ctcctccaac | tccagcagct | cttttcccc | cacagccagt | 2640 |
| gaacaacgcg | gacttcatca | tccctgtgga | gattgatggg | gtggtccacc | aggtgagggc | 2700 |
| caggaagagg | cagtggtggg | cttggcatct | gcctccagac | cctaggctct | tcccaccaat | 2760 |
| ccggagcgcc | tcggatggga | attggataca | tgtggaatgt | cagaggccca | gagaggggtgt | 2820 |
| gagacttgtc | ccaaagtac | acagaacctc | aagggcttgt | gctgactcca | agcctgcaga | 2880 |
| gtgggctcct | cctctaggct | ccccctgtct | gtgctccctc | gccccaccct | gcccgggacc | 2940 |
| cagttcaagt | aattcaggat | aggttgtgt | ctgtccagcc | tgttctccat | tacttggctc | 3000 |
| ggggaccgg | gccctgcagc | cttggggta | gggggctgccc | cctggattcc | tgcactaggc | 3060 |
| tgaggttag | gcaggggaag | ggattggaa | ttagggacct | cgtgaggtag | gactggccag | 3120 |
| tggagttggaa | gttttgc | ttttctggcg | gggggtgggt | acagttccc | cagcagtgg | 3180 |
| caggtagct | ggccaagcgg | agcctgcggg | cccagtctcc | ttcctgtgcg | cctctgcctc | 3240 |
| cctggcccat | gccctgccag | ccctcgccca | cccccacact | gccccactgg | cccgagccc | 3300 |
| cctcaactggc | cgcccccca | ggtctacgt | ttgaagcgtc | ctcatgtgga | ttagttcctg | 3360 |
| cagcgaatgg | gcgagcttt | tgaatgtgt | ctgttcactg | ctagcctcgc | caaggtgagc | 3420 |
| cccacagggg | tcccggggca | accctgcct | cctacctacc | tcccgcac | agcccagtga | 3480 |
| acctgcgggc | cccaggatga | cccacccct | gctcccagta | cgcagaccca | gtagctgacc | 3540 |
| tgctggacaa | atggggggcc | ttccggggcc | ggctgtttcg | agagtcctgc | gtttccacc | 3600 |
| gggggaacta | cgtgaaggac | ctgagccggt | tgggtcgaga | cctgcggcgg | gtgctcatcc | 3660 |
| tggacaattc | acctgcctcc | tatgtttcc | atccagacaa | tgctgtgagt | gcgggctgga | 3720 |
| ctgggactgg | gacaggagct | gagaccagg | aagggtcag | tccattcagg | ccaccttggc | 3780 |
| ctcttggatc | cccagttggg | gggtgggtgc | cctcccagtc | ttcctgcac | tcattgcctg | 3840 |
| tgcctgccgc | ccactccct | catccacctg | ccctgttagcc | atatggtctt | ttcccctcgc | 3900 |
| acaaagcaga | gcatctgcca | tgcacagggg | cccccacagg | gcaacggagt | ttggaaagtt | 3960 |
| tcaattttc | gaattgccag | ttgtgaccta | ctgatggccc | acagaattaa | tttagtgggt | 4020 |
| tctgattggg | aatttaaca | aaatgaaata | aatagaaaa | tatccggcgt | ggtgcagtgg | 4080 |

tcatgcctg taatcccagc actttggaa gctgagggtgg gcaggttagct gagcccagta 4140
gttcaagacc agcctcgca acatagtcaa accttatgtc tacaaaaaat acaaaaacta 4200
gccaggcgtg gtggcgcatg cctggagtcc cggttatgca gaaggctgag gtaggaggtat 4260
cgcttgagcc ctggaggcag aggctgtggt gagccaagat tgtgccactg cactctagcc 4320
tggcaacag agcaagaccc tgccctcaaaa aaaaaaaaaa gtatccaagt gcttcgcaca 4380
gataaggta ggaattgtga agctttgca ttgttacgtt ataaatgtgt tttcctgggg 4440
attgctgtca aaaaagttt aacactgtgg gtgaggggtt ttcagaaact gcatgatctg 4500
agtagtggtc acataggcgt ggcctggaaa ttctgcaccc aggaccacct gccccctca 4560
tcttcctaca cccacttccc caggtaccgg tggcctcgat gtttgcacaat atgagtgaca 4620
cagagctcca cgaccccttc cccttctcg agcaactcag ccgtgtggac gacgtgtact 4680
cagtgcctcag gcagccacgg ccagggagct agtgagggtg atggggccag gacctgcccc 4740
tgaccaatga tacccacacc tcctcccagg aagactgccc aggcctttgt tagaaaaacc 4800
catggccgcg cggcacactc agtgccatgg ggaagcgggc gtctcccca ccagccccac 4860
caggcggtgt agggcagca ggctgcactg aggaccgtga gctccaggcc ccgtgtcagt 4920
gccttcaaacc ctccctccct attctcaggg gacctgggg gccctgcctg ctgctccctt 4980
tttctgtctc tgccatgct gccatgttcc tctgctgcca aattggggccc cttggccccct 5040
tccggttctg ctccctgggg gcaggggtcc tgccttggac ccccagtctg ggaacgggtgg 5100
acatcaagtgcctgcata agccccctct tccccgcccc gcttcccag gggcacagct 5160
ctaggctggg aggggagaac cagccccctcc ccctgccccca cctccctccct tgggactgag 5220
agggcccccta ccaaccttg cctctgcctt ggagggaggg gaggtctgtt accactgggg 5280
aaggcagcag gagtctgtcc ttcaggcccc acagtgcagc ttctccaggcc cgacagctg 5340
agggctgctc cctgcatacat ccaagcaatg acctcagact tctgccttaa ccagccccgg 5400
ggcttggctc cccagctct gagcgtgggg gcataggcag gaccccccctt gtggtgccat 5460
ataaaatatgt acatgtgtat atagatttt agggaaagga gagagggaaag ggtcagggtat 5520
gagacaccccc tcccttgccc ctttcctggg cccagaagtt ggggggaggg agggaaagga 5580
tttttacatt ttttaaactg ctatttctg aatggaacaa gctggggccaa gggggccagg 5640
ccctgtcctc tgccatcacc accccttgc tccggttcat cattcaaaaa aacatttctt 5700
gagcacccctc tggcccccagc atatgctagg cccaccagct aagtgtgtgt ggggggtctc 5760
tacgccagct catcagtgcc tccttgcccc tccttcaccg gtgcctttgg gggatctgtat 5820
ggaggtggga cttctgtgg ggtttggggta tctccaggaa gcccggccaa gctgtccccct 5880

| | | | | | | |
|-------------|-------------|-------------|-------------|------------|-------------|------|
| tcccctgtgc | caaccatct | cctacagccc | cctgcctgat | cccctgctgg | ctgggggcag | 5940 |
| ctcccaggat | atcctgcctt | ccaaactgttt | ctgaagcccc | tcctctaacc | atggcgattc | 6000 |
| cggaggtcaa | ggccttggc | tctcccagg | gtctaacgg | taaggggacc | cacataccag | 6060 |
| tgccaagggg | gatgtcaagt | ggtgatgtcg | ttgtgctccc | ctccccaga | gcgggtggc | 6120 |
| gggggggtgaa | tatggttggc | ctgcatcagg | tggccttccc | attnaagtgc | cttctctgtg | 6180 |
| actgagagcc | ctagtgtgat | gagaactaaa | gagaaagcca | gaccctatac | ctgcttctgt | 6240 |
| ggttattgcg | ggggacttca | gcaagtgggg | tgtgtgcctt | gcacctgcgg | ctgccgtgg | 6300 |
| cccccccccc | gcttcagcac | acctagaggg | ctgttggtgg | agggaggggc | tgcccgcccc | 6360 |
| tcgacacttc | aggtgggaag | ggcagcgtca | gagcacaaat | ttgagcctcc | aggctgtgct | 6420 |
| cgtctacgtc | ttcccgccctc | gggtatgtgg | tctgc当地at | ggagatgtgc | cctattggca | 6480 |
| ggactaatta | agtgcctgga | cacagacgac | aggatactag | tagctggaaa | gcaaaattcg | 6540 |
| aaggcctggg | taggggcagt | cctggaatgc | ggcggggggag | ggggcgtggc | ctctgccc | 6600 |
| gagcagaggg | gcggggcttg | tgcggctccg | aaggcagagg | cggggagcgg | ggcgaggctc | 6660 |
| tgggtggagg | ctccagcggc | agaacttgg | ggcctgggtg | cggcgggctc | cggcgcctgg | 6720 |
| ctctgccggg | cggcctgggt | ggggccggcg | ccggggctcg | cccccccccg | cccctctgcg | 6780 |
| gcctctgagc | agccattggc | cgcgc当地cc | ccccacttcc | cgc当地ccccc | cgcgtccggg | 6840 |
| aggcacttcc | tttgc当地aaac | cgc当地ccccc | caggcgc当地gg | caggaaatgc | cctccgc当地gg | 6900 |
| tccccagcca | gccttgctt | gcttcccacg | ccagccgcta | gaggcctccc | tgtcctcg | 6960 |
| gacgcaggaa | ctccccgggg | gctggaaaga | tggggccac | ctcactcacc | cccttcccgg | 7020 |

<210> 14
 <211> 4833
 <212> DNA
 <213> Homo sapiens

| | | | | | | |
|-------------|------------|------------|------------|------------|-------------|-----|
| <400> 14 | 60 | | | | | |
| gccatttcct | cctttgttt | tcactccgga | ttctccatgt | tggacccaaa | ctgaggagcc | 60 |
| cggagctgcc | gctgggggat | cggggccggg | ggcacccggg | ggagccgctg | cccgccgc | 120 |
| ccgc当地tttgc | tacaggccgc | ctcccttccc | ggtccgggaa | ggaaacgaga | ggggggatgt | 180 |
| gaacagctgt | ggaagtgc当地 | gtctcggag | ccggagcggg | ccccgc当地ca | ggccccccag | 240 |
| cccagccag | cccgc当地cc | cgc当地ccct | cccgtccagc | cagcccggc | ccgc当地ggatt | 300 |
| gttagatgga | acacggctcc | atcatcaccc | aggcgc当地gg | ggaagacg | cc | 360 |
| ccaagcaagg | cctggctcc | aagtccctc | ctaagaagcc | tcgtggacgt | aacatcttca | 420 |
| aggccctttt | ctgctgtttt | cgc当地ccagc | atgttggcca | gtcaagttcc | tccactgagc | 480 |

| | | | | | | |
|------------|-------------|-------------|-------------|-------------|-------------|------|
| tcgctgcgt | taaggaggaa | gcaaacacca | ttgctaagtc | ggatctgctc | cagtgtctcc | 540 |
| agtaccagtt | ctaccagatc | ccagggacct | gcctgctccc | agaggtgaca | gaggaagatc | 600 |
| aaggaaggat | ctgtgtggtc | attgacactcg | atgaaaccct | tgtgcata | tcctttaagc | 660 |
| caatcaacaa | tgctgacttc | atagtgccta | tagagattga | ggggaccact | caccaggtgt | 720 |
| atgtgctcaa | gaggcattat | gtggatgagt | tcctgagacg | catggggaa | ctcttgaat | 780 |
| gtgttctt | cactgccagc | ctggccaagt | atgcccaccc | tgtgacagac | ctgctggacc | 840 |
| ggtgtgggt | gttccgggccc | cgcctattcc | gtgagtcttg | cgtgttccac | cagggctgct | 900 |
| acgtcaagga | cctcagccgc | ctggggaggg | acctgagaaa | gaccctcatc | ctggacaact | 960 |
| cgcctgcttc | ttacatattc | caccccgaga | atgcagtgcc | tgtgcagtcc | tggttttagt | 1020 |
| acatggcaga | cactgagttg | ctgaacctga | tcccaatctt | tgaggagctg | agcggagcag | 1080 |
| aggacgtcta | caccagcctt | ggggcagctg | cggggccctt | agcctgcct | gcttccaagc | 1140 |
| gacggccatc | ccagtagggg | actttccac | actgtgcctt | tacgatcagc | gtgacagagt | 1200 |
| agaagctgga | gtgcctcacc | acacggcccg | gaaacagcgg | gaagtaactg | gaaagagctt | 1260 |
| taggacagct | tagatgccga | gtggcgaat | gccagaccaa | tgataacccag | agctacactgc | 1320 |
| cgc当地actt | ttgagatgtg | tgtttactg | ttagagagtg | tgtgtttgtg | tgtgtgtttt | 1380 |
| gccatgaact | gtggccccag | tgtatagtgt | ttcagtgggg | gagaagctga | aagaccaaga | 1440 |
| ctcttcccaa | gttagcttgt | ctcctctcct | gtcaccctaa | gagccactga | gttgtgttagg | 1500 |
| gatgaaract | attgaagact | ccattgccaa | accatggcct | ttcctcagtg | ttgtaaggcc | 1560 |
| tatgccaagg | ataaaggaag | ggtatgcctt | tgggtactcc | aggcatacac | ctttctgaaa | 1620 |
| tccttctcca | gccagctgct | gcagacaaaa | gatcacattt | ctgggaagat | gagaacttgt | 1680 |
| ttccagacca | gcatccagtg | gccatcaggt | cttgtggccc | aaaggctatg | ctgcctccg | 1740 |
| gctgagtgcc | tgggataggc | cttttctatg | tctcccaag | gctgggggtgc | tgagcctgccc | 1800 |
| ttcctcacca | cctagccata | gtctcaaacc | tgtggggaaag | gaggtttct | ccctgcccgg | 1860 |
| gaagaggaca | gataactgat | ttccgttctt | ttgactgtgt | tttaaaattc | tctttctaaa | 1920 |
| cacagagtgt | tgggcctgg | ttgtttctga | caaagttaca | gtcctgggcc | tgtaatgaat | 1980 |
| gtcggcggcg | ctggggttgc | agggaaaaga | caaatcctca | aagcgtggac | gtgtgtcccc | 2040 |
| atggcttgc | gatcagctaa | gctcgggatc | attccataaa | gtctgtttt | cagggattct | 2100 |
| ctgctggtgc | tggtgcaagg | acttctgttc | caaaggctgg | aaaaaactaa | gctgtcccag | 2160 |
| cccctccat | ttcttgggca | gggcttttt | cctgttgtgt | cttccccag | ggcctgtcct | 2220 |
| gtaccgagct | ctgtctgttc | cagcctacat | cttccctggg | tgttgctttt | cctcttaagg | 2280 |
| gcctcagaac | tcttgctctt | cctggggtga | ggggaaatga | gtgttcttga | catgtgacag | 2340 |

| | |
|---|------|
| ccta atgcgc atgcttctg cctctggtaa caggagttag tgagccctc agacctgcac | 2400 |
| tctgggtgtc tcctgcttac aaaggttctt aatagtgaat gctttaaaat taaagtcatc | 2460 |
| acgaaatgga agtttccca gggtggaaaa taagaggaag tgctgctgta attgggagca | 2520 |
| caaggggcct cccaaaaagg agccccacct cagcatcaact gccttaatcg tggcctccct | 2580 |
| ggggtggtgt gggttctctc ctcccctccct ccctcctccct ggggtggag ggcgctcctg | 2640 |
| ttcccatctc tgtgttccct ggaggcaggt atcacaaagc atttgtgaat tgcttaggt | 2700 |
| gcagggacac cacccactca ggactttcc ccatcatccc ttccattgcc acaccctaga | 2760 |
| tccagcctca ggaactaaca agttktgaga aaagcaggtg gttagcagc agcttcgtgc | 2820 |
| tctcagcgggt ggctggctgg cattttctc tagcgttgtg gtgccacctt cccttcttgt | 2880 |
| cccaaggta taaggccttg tcttctctt tggaatcata aagtggaca gagtccccag | 2940 |
| aactcatgtg ghcatttccg acagcatcac tccccgggtgc ctatggggtc ccgggtgtacc | 3000 |
| taaaggaga aggacccat gtgctagcca gaaataact gtctctgaa ggaaagcagg | 3060 |
| agctcagact cttagagcca gctgtggctt cggacccaag gcctgaccta ggctgctatc | 3120 |
| ctaattttgg aggaggggcc tctcttccaa gccccaccct aagggttagc cttggacaa | 3180 |
| atcttgcc gtc taggccc agccaggctt ttctgactaa ataagcaata agaggctcta | 3240 |
| agctgactga gttgcaagga cccttccgc cctcccttgg atctccatgt ttctccagat | 3300 |
| ggcggaaagag catgtgccac ccccttcct aacagacttg tccaagtgt tgccgtggaa | 3360 |
| cccatgacca aagcccagga tggctggtg ggagtgtccc tgctgcacat gcatgaagcc | 3420 |
| cctgctttt aggctcaact cccatcagaa ccctgcctgc ccacctgcaa ctccccccca | 3480 |
| acaatgccat tcccacttgc cccagagaag ctactcggcc aaacctagcc agggctgtt | 3540 |
| cttggacc agagccagcc tagtattat ttgctgtcg gttccagtt tcaccgtgt | 3600 |
| ttagggtgag ggttgcattgt aaaatttgct cctcaaagga atcaggccag actcaatttt | 3660 |
| gggagggcaa gacagggagg aggccgcttc atcccagact ctcttctagg gcttcccacc | 3720 |
| atcagccctt cccacttgc actggctttt gggaggcaat aggccaccat gcctggtcag | 3780 |
| caccaattca agccatgccca ggaatgtgcc tacctgccag gttcagttct tttaggtgc | 3840 |
| ctcttcaggg acacagtgtg tctctctgat tggcttcta aatcaaaaagc ctgatgtcg | 3900 |
| tgtccctctc atagggggag cttggacac aggaccagtt tgaaaaagg tcaggttaagg | 3960 |
| gtttccactc tgcacattgt agagggaca ctctgttaggc ccatgggtcc cttaactagag | 4020 |
| aggtttagtg aatttgccct cagtaacat gggacccctt gtttagcttc ctcttgcttc | 4080 |
| ccaaagattt taagcatttt gttaaatgtat aaactcacct ctggtaaacag tggcccagac | 4140 |

| | | | | | | |
|------------|------------|------------|------------|------------|-------------|------|
| gctgcttgt | gctaaaagca | tggaaatgt | aaaggcagtc | tttctctgg | aatggatgc | 4200 |
| tattctattc | tgctgccct | acctgttcct | gaggcctcat | ttagaaagaa | aatcccctca | 4260 |
| gaaggctgtc | tggcacccag | tgtcctagcc | aggccaagta | tatgagaaag | gtaaagtccat | 4320 |
| tttcccttc | aggtcctcag | tggattactt | aaccactgct | gtccctcggt | cccttttcc | 4380 |
| taaacgggtt | tagttctgtc | tttttctcc | tttttctaa | atgctggtaa | atatttacat | 4440 |
| tcagccaggg | aagaggaggc | cagaggtcgg | gccagctgcc | ccatttttt | aacgttntag | 4500 |
| ggcctgccc | tggagcggac | cctcctctt | gggcctcgtg | agctttttg | cttatcatgt | 4560 |
| tccatttcgt | gccgcttcc | cccttcaaga | tgccatttg | aggtagggg | atctgcttcc | 4620 |
| cactgtgact | gggctatggg | attctgacta | ccttgcttac | agattcatgg | tttgataaat | 4680 |
| ttgttgtatt | ccaaaacttg | aatgcagga | cgccattaag | tgtctgtta | tatTTTgga | 4740 |
| atatttgtat | tacttacaat | taattaataa | aagtgggtt | aaaaaacctt | tccaggaaaa | 4800 |
| aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | aaaaaaa | aaa | | 4833 |

<210> 15
<211> 859
<212> DNA
<213> Homo sapiens

| | | | | | | |
|-------------|-------------|------------|------------|------------|------------|-----|
| <400> 15 | | | | | | |
| atggacggcc | cggccatcat | cacccaggtg | accaacccca | aggaggacga | gggcccgttg | 60 |
| ccgggcgcgg | gcgagaaagc | ctcccagtgc | aacgtcagct | taaagaagca | gaggagccgc | 120 |
| agcatccta | gctccttctt | ctgctgcttc | cgtgattaca | atgtggaggc | ccctccaccc | 180 |
| agcagccccca | gtgtgcttcc | gccactggtg | gaggagaatg | gtgggcttca | gaagccacca | 240 |
| gctaagtacc | ttcttccaga | ggtgacggtg | cttgactatg | gaaagaaatg | tgtggtcatt | 300 |
| gattnagatg | aaacattgg | gcacagttcg | tttaagccta | ttagtaatgc | tgattttatt | 360 |
| gttcccggtt | aatcgtatgg | aactatacat | caggtgtatg | tgctgaagcg | gccacatgtg | 420 |
| gacgagttcc | tccagaggat | ggggcagctt | tttgaatgtg | tgctctttac | tgccagcttg | 480 |
| gccaagtatg | cagaccctgt | ggctgacctc | ctagaccgct | ggggtgtgtt | ccggggcccg | 540 |
| ctcttcagag | aatcatgtgt | tttcatcgt | gggaactacg | tgaaggacct | gagtcgcctt | 600 |
| gggcgggagc | tgagcaaagt | gatcattgtt | gacaattccc | ctgcctcata | catcttccat | 660 |
| cctgagaatg | cagtgcctgt | gcagtcctgg | ttcgatgaca | tgacggacac | ggagctgctg | 720 |
| gacctcatcc | cttctttga | gggcctgagc | cgggaggacg | acgtgtacag | catgctgcac | 780 |
| agactctgca | ataggttagcc | ctggcctctg | cctgcctccc | gcctgtgcac | tctggaacct | 840 |
| ctggcctcag | gggacctgc | | | | | 859 |

| | | | | | | |
|--------------|--|------------|--------------|-------------|-------------|-----|
| <210> | 16 | | | | | |
| <211> | 754 | | | | | |
| <212> | DNA | | | | | |
| <213> | Homo sapiens | | | | | |
| | | | | | | |
| <400> | 16 | | | | | |
| atgatgggga | ggccctgcct | gctcacagcg | gggcgc(cc)ct | gcttgtggag | gagaatggcg | 60 |
| ccatccc(taa) | ggcagacccc | agtccaatac | ctgctccctg | aggccaaggc | ccaggactca | 120 |
| gacaagatct | gcgtggtcat | cgacctggac | gagaccctgg | tgcacagctc | cttcaagcca | 180 |
| gtgaacaacq | cggacttcat | catccctgtg | gagattgatg | gggtggtcca | ccaggtctac | 240 |
| gtgttgaagc | gtcctcacgt | ggatgagttc | ctgcagcgaa | tggcgagct | ctttgaatgt | 300 |
| tgctgttca | ctgcttagcct | cgc(a)gtac | gcagacccag | tagctgac(t) | gctggacaaa | 360 |
| tggggggcct | tccgggccc | gctgttcga | gagtcc(t)cg | tcttccaccg | ggggactac | 420 |
| gtgaaggacc | t(g)agccggtt | gggtcgagac | ctgcggcggg | tgctcatcct | ggacaattca | 480 |
| cctgcctcct | atgtttcca | tccagacaat | gctgtaccgg | tggcctcg(t) | gtttgacaac | 540 |
| atgagtgaca | cagagctcca | cgacccctc | cccttctcg | agcaactcag | ccgtgtggac | 600 |
| gacgtgtact | cagtgctcag | gcagccacgg | ccagggagct | agtgagggtg | atggggccag | 660 |
| gacctgcccc | tgaccaatga | tacccacacc | tcctcccagg | aagactgccc | aggc(t)ttgt | 720 |
| tagaaaaacc | catgggcccgc | cgcacactc | agt | | | 754 |
| | | | | | | |
| <210> | 17 | | | | | |
| <211> | 27 | | | | | |
| <212> | DNA | | | | | |
| <213> | Artificial sequence | | | | | |
| | | | | | | |
| <220> | | | | | | |
| <223> | Synthetic construct: polymerase binding site | | | | | |
| | | | | | | |
| <400> | 17 | | | | | |
| gaattaaatac | gactcactat | agggaga | 27 | | | |
| | | | | | | |
| <210> | 18 | | | | | |
| <211> | 25 | | | | | |
| <212> | DNA | | | | | |
| <213> | Artificial sequence | | | | | |
| | | | | | | |
| <220> | | | | | | |
| <223> | Primer | | | | | |
| | | | | | | |
| <400> | 18 | | | | | |
| atgggcgaac | tatacgagt | cgttc | 25 | | | |
| | | | | | | |
| <210> | 19 | | | | | |
| <211> | 25 | | | | | |
| <212> | DNA | | | | | |

| | |
|-----------------------------|----|
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 19 | 25 |
| atcaacgaca acttcgagat cgtcg | |
| <210> 20 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 20 | 24 |
| atgtcgctct tgcaaaaaact aagc | |
| <210> 21 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 21 | 25 |
| tgaagatcct caccgagcgc ggcta | |
| <210> 22 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 22 | 25 |
| cagctggtgc gggagtacgg cttcc | |
| <210> 23 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 23 | 22 |
| gagctgtcgt tgagctttgg cg | |
| <210> 24 | |
| <211> 23 | |
| <212> DNA | |
| <213> Artificial sequence | |

| | | |
|------------------------------|--|----|
| <220> | | |
| <223> Primer | | |
| <400> 24 | | |
| actgggccta ttactactgg ctc | | 23 |
| <210> 25 | | |
| <211> 25 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 25 | | |
| caacgaagcc gagcgagcca tccag | | 25 |
| <210> 26 | | |
| <211> 26 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 26 | | |
| gcaacaactg gccaagggt cattac | | 26 |
| <210> 27 | | |
| <211> 26 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 27 | | |
| gccttccaag agcacgacgt acaaag | | 26 |
| <210> 28 | | |
| <211> 26 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 28 | | |
| ctcgccaatc aagtaccttg tgctgc | | 26 |
| <210> 29 | | |
| <211> 25 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |

| | | |
|------------------------------|--|----|
| <220> | | |
| <223> Primer | | |
| <400> 29 | | |
| cttcgctcgc acctcagaaa cgatc | | 25 |
| <210> 30 | | |
| <211> 25 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 30 | | |
| caacggaaact aacggccgct ccgag | | 25 |
| <210> 31 | | |
| <211> 22 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 31 | | |
| ctcgccattg ttctcctgggt gg | | 22 |
| <210> 32 | | |
| <211> 25 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 32 | | |
| cttgtctgct gctggttcaa catgg | | 25 |
| <210> 33 | | |
| <211> 25 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 33 | | |
| gcgggttggag tagccaaact cgttg | | 25 |
| <210> 34 | | |
| <211> 25 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |

| | |
|------------------------------|----|
| <223> Primer | |
| <400> 34 | |
| ttataggata tcttcgattt tcggc | 25 |
| <210> 35 | |
| <211> 26 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 35 | |
| gaccggactc gtcatactcc tgcttg | 26 |
| <210> 36 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 36 | |
| tgcgcagctc gcccatgttag acctg | 25 |
| <210> 37 | |
| <211> 23 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 37 | |
| cgcgtggatt gggagaagaag gtc | 23 |
| <210> 38 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 38 | |
| ccgtaaaacc gcgcgcatta aagt | 24 |
| <210> 39 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |

| | |
|-------------------------------|----|
| <400> 39 | |
| tggcatggt cacgaatccg aatc | 24 |
| | |
| <210> 40 | |
| <211> 27 | |
| <212> DNA | |
| <213> Artificial sequence | |
| | |
| <220> | |
| <223> Primer | |
| | |
| <400> 40 | |
| cttggcatcg aacatctgct gggtaag | 27 |
| | |
| <210> 41 | |
| <211> 26 | |
| <212> DNA | |
| <213> Artificial sequence | |
| | |
| <220> | |
| <223> Primer | |
| | |
| <400> 41 | |
| cgatcagaag tggatcgccgg tcctta | 26 |
| | |
| <210> 42 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial sequence | |
| | |
| <220> | |
| <223> Primer | |
| | |
| <400> 42 | |
| ccctggctga agcagaactt catg | 24 |
| | |
| <210> 43 | |
| <211> 25 | |
| <212> DNA | |
| <213> Artificial sequence | |
| | |
| <220> | |
| <223> Primer | |
| | |
| <400> 43 | |
| tatggcataa aagggtggc cattc | 25 |
| | |
| <210> 44 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial sequence | |
| | |
| <220> | |
| <223> Primer | |

| | |
|-------------------------------|----|
| <400> 44 | 24 |
| gttctcgcca tcggtgagat ctgc | |
| | |
| <210> 45 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial sequence | |
| | |
| <220> | |
| <223> Primer | |
| | |
| <400> 45 | 22 |
| cgtacatgag gtagaccctg ga | |
| | |
| <210> 46 | |
| <211> 27 | |
| <212> DNA | |
| <213> Artificial sequence | |
| | |
| <220> | |
| <223> Primer | |
| | |
| <400> 46 | 27 |
| cggccgtcat tactcagatc agcaagg | |
| | |
| <210> 47 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial sequence | |
| | |
| <220> | |
| <223> Primer | |
| | |
| <400> 47 | 22 |
| tccaccaccc tgtgttgctg ta | |
| | |
| <210> 48 | |
| <211> 27 | |
| <212> DNA | |
| <213> Artificial sequence | |
| | |
| <220> | |
| <223> Primer | |
| | |
| <400> 48 | 27 |
| catctctgat ctcgactgct ccagcag | |
| | |
| <210> 49 | |
| <211> 30 | |
| <212> DNA | |
| <213> Artificial sequence | |
| | |
| <220> | |
| <223> Primer | |
| | |
| <400> 49 | |

| | |
|----------------------------------|----|
| tgcctcacc caaggctct gacactgtgg | 30 |
| <210> 50 | |
| <211> 30 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 50 | |
| ctgtggccat ggaggaaac agtggttcc | 30 |
| <210> 51 | |
| <211> 28 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 51 | |
| gcaaccgcag gcacgactgt ttacggag | 28 |
| <210> 52 | |
| <211> 29 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 52 | |
| ccatcgccctg cgaaacctcc ccaggtaga | 29 |
| <210> 53 | |
| <211> 28 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 53 | |
| gcagtgaaca gcacacattc aaagagct | 28 |
| <210> 54 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial sequence | |
| <220> | |
| <223> Primer | |
| <400> 54 | |
| accacagtc accatcac | 20 |

| | | |
|----------------------------------|--|----|
| <210> 55 | | |
| <211> 27 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 55 | | |
| gggtcagaga gtggatgc cacagtg | | 27 |
| <210> 56 | | |
| <211> 28 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 56 | | |
| cttgaacagc tcctggatgg cagtgtg | | 28 |
| <210> 57 | | |
| <211> 28 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 57 | | |
| agaagtccag gagcagctga gggagcac | | 28 |
| <210> 58 | | |
| <211> 30 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 58 | | |
| agatgaccat ccggaagaag ttggcattgt | | 30 |
| <210> 59 | | |
| <211> 30 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| <220> | | |
| <223> Primer | | |
| <400> 59 | | |
| agccaactca gctggactct ctccagcttc | | 30 |

| | | |
|-----------------------------------|--|----|
| <210> 60 | | |
| <211> 30 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| | | |
| <220> | | |
| <223> Primer | | |
| | | |
| <400> 60 | | |
| tgcggttat attatcctgc acgcccggag | | 30 |
| | | |
| <210> 61 | | |
| <211> 27 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| | | |
| <220> | | |
| <223> Primer | | |
| | | |
| <400> 61 | | |
| ggagccctat gcagggttaag ggaataa | | 27 |
| | | |
| <210> 62 | | |
| <211> 28 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| | | |
| <220> | | |
| <223> Primer | | |
| | | |
| <400> 62 | | |
| aactatttct gggtaactcc ttagacac | | 28 |
| | | |
| <210> 63 | | |
| <211> 30 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| | | |
| <220> | | |
| <223> Primer | | |
| | | |
| <400> 63 | | |
| ctggataagt tactgaagag tgggctttgg | | 30 |
| | | |
| <210> 64 | | |
| <211> 29 | | |
| <212> DNA | | |
| <213> Artificial sequence | | |
| | | |
| <220> | | |
| <223> Primer | | |
| | | |
| <400> 64 | | |
| caccgggttcg agtccccgga gaggatatac | | 29 |

```

<210> 65
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 65
ggccttgat tttggagcc accttgtg 28

<210> 66
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 66
gctgggagga atgcttcta atgcatttg 29

<210> 67
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 67
cagacgacaa gttacatgca acatg 25

<210> 68
<211> 150
<212> PRT
<213> Homo Sapiens

<400> 68

Asn Arg Lys Leu Val Leu Met Val Asp Leu Asp Gln Thr Leu Ile His
1 5 10 15

Thr Thr Glu Gln His Cys Gln Gln Met Ser Asn Lys Gly Ile Phe His
20 25 30

Phe Gln Leu Gly Arg Gly Glu Pro Met Leu His Thr Arg Leu Arg Pro
35 40 45

His Cys Lys Asp Phe Leu Glu Lys Ile Ala Lys Leu Tyr Glu Leu His
50 55 60

Val Phe Thr Phe Gly Ser Arg Leu Tyr Ala His Thr Ile Ala Gly Phe
65 70 75 80

```

Leu Asp Pro Glu Lys Lys Leu Phe Ser His Arg Ile Leu Ser Arg Asp
85 90 95

Glu Cys Ile Asp Pro Phe Ser Lys Thr Gly Asn Leu Arg Asn Leu Phe
100 105 110

Pro Cys Gly Asp Ser Met Val Cys Ile Ile Asp Asp Arg Glu Asp Val
115 120 125

Trp Lys Phe Ala Pro Asn Leu Ile Thr Val Lys Lys Tyr Val Tyr Phe
130 135 140

Gln Gly Thr Gly Asp Met
145 150

<210> 69

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Consensus sequence

<220>

<221> MISC_FEATURE

<222> (1)..(3)

<223> Xaa at 1, 2, or 3 is a hydrophobic residue

<220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> Xaa is any amino acid

<220>

<221> MISC_FEATURE

<222> (7)..(7)

<223> Xaa is any amino acid

<220>

<221> MISC_FEATURE

<222> (8)..(8)

<223> Xaa is T or V

<220>

<221> MISC_FEATURE

<222> (9)..(10)

<223> Xaa is a hydrophobic amino acid

<400> 69

Xaa Xaa Xaa Asp Xaa Asp Xaa Xaa Xaa Xaa
1 5 10